

# Xuanjun Chen

✉ d12942018@ntu.edu.tw | 🏠 xjchen.tech | 📄 xjchenGit | 📺 jun-ntu | 🎓 Google Scholar

## Education

### National Taiwan University (NTU)

PH.D. STUDENT IN GRADUATE INSTITUTE OF COMMUNICATION ENGINEERING, GPA 4.30/4.30

Taipei Taiwan

Sep. 2023 - Present

M.S. IN COMPUTER SCIENCE AND INFORMATION ENGINEERING, GPA 4.19/4.30

Sep. 2020 - Jan. 2023

- **Advisor:** Prof. [Jyh-Shing Roger Jang](#) and Prof. [Hung-yi Lee](#)

### National Taiwan University of Science and Technology (NTUST)

B.S. IN COMPUTER SCIENCE AND INFORMATION ENGINEERING, GPA 4.11/4.30

Taipei Taiwan

Sep. 2018 - Jun. 2020

### Shenzhen Univ. of Information Technology (f.k.a. Shenzhen Institute of Information Technology, SZIT)

Shenzhen China

JUNIOR COLLEGE IN ELECTRONIC INFORMATION ENGINEERING TECHNOLOGY

Sep. 2015 - Jul. 2018

## Selected Publications († indicates equal contribution.)

- [1] **Xuanjun Chen**<sup>†</sup>, Jiawei Du<sup>†</sup>, Haibin Wu, Lin Zhang, I-Ming Lin, I-Hsiang Chiu, Wenze Ren, Yuan Tseng, Yu Tsao, Jyh-Shing Roger Jang, Hung-yi Lee, “[CodecFake-Omni: A Large-Scale Codec-based Deepfake Speech Dataset](#)” *Submitted to TASLP, 2025.*
- [2] Chien-yu Huang, Wei-Chih Chen, Shu-wen Yang, Andy T. Liu, Chen-An Li, Yu-Xiang Lin, Wei-Cheng Tseng, Anuj Diwan, Yi-Jen Shih, Jiatong Shi, William Chen, **Xuanjun Chen**, et al., “[Dynamic-SUPERB Phase-2: A Collaboratively Expanding Benchmark for Measuring the Capabilities of Spoken Language Models with 180 Tasks](#)”. in *ICLR, Singapore, April 2025.*
- [3] Wenze Ren, Haibin Wu, Yi-Cheng Lin, **Xuanjun Chen**, Rong Chao, Kuo-Hsuan Hung, et al., “[Leveraging Joint Spectral and Spatial Learning with MAMBA for Multichannel Speech Enhancement](#).” in *ICASSP 2025, Hyderabad, India, April 2025.*
- [4] **Xuanjun Chen**, Haibin Wu, Chung-Che Wang, Hung-yi Lee, Jyh-Shing Roger Jang “[Multimodal Transformer Distillation for Audio-Visual Synchronization](#).” in *ICASSP 2024, Seoul, Korea, April 2024.*
- [5] **Xuanjun Chen**, Haibin Wu, Jyh-Shing Roger Jang, Hung-yi Lee “[Singing Voice Graph Modeling for SingFake Detection](#).” in *Interspeech 2024, Kos Island, Greece, Sept 2024.*
- [6] **Xuanjun Chen**<sup>†</sup>, Jiawei Du<sup>†</sup>, Haibin Wu, Jyh-Shing Roger Jang, Hung-yi Lee “[Neural Codec-based Adversarial Sample Detection for Speaker Verification](#)”. in *Interspeech 2024, Kos Island, Greece, Sept 2024.*
- [7] Haibin Wu, **Xuanjun Chen**, Yi-Cheng Lin, Kai-wei Chang, Ho-Lam Chung, Alexander H. Liu, Hung-yi Lee “[Towards audio language modeling-an overview](#).” *Overview Report, Feb. 2024.*
- [8] Haibin Wu, Ho-Lam Chung, Yi-Cheng Lin, Yuan-Kuei Wu, **Xuanjun Chen**, Yu-Chi Pai, et al., “[Codec-SUPERB: An In-Depth Analysis of Sound Codec Models](#)”. in *Findings of ACL 2024, Bangkok, Thailand, Aug 2024.*
- [9] Haibin Wu, **Xuanjun Chen**, Yi-Cheng Lin, et al., “[Codec-SUPERB @ SLT 2024: A lightweight benchmark for neural codec models](#)”. in *IEEE SLT 2024, Macao, China, 2024.*
- [10] Jiawei Du, I-Ming Lin, I-Hsiang Chiu, **Xuanjun Chen**, Haibin Wu, Wenze Ren, Yu Tsao, Hung-yi Lee, Jyh-Shing Roger Jang “[DFADD: The Diffusion and Flow-Matching based Audio Deepfake Dataset](#).” in *IEEE SLT 2024, Macao, China, 2024.*
- [11] Hsuan-Yu Lin, **Xuanjun Chen**, Jyh-Shing Roger Jang “[Singer separation for karaoke content generation](#).” in *The 27th Oriental COCOSDA Conference (O-COCOSDA 2024), Hsinchu, Taiwan, Oct 2024.*
- [12] **Xuanjun Chen**<sup>†</sup>, Haibin Wu<sup>†</sup>, Helen Meng, Hung-yi Lee, Jyh-Shing Roger Jang “[Push-Pull: Characterizing the Adversarial Robustness for Audio-Visual Active Speaker Detection](#).” in *IEEE SLT 2022, Doha, Qatar, Jan 2023.*
- [13] **Xuanjun Chen**<sup>†</sup>, Yen-Lun Liao<sup>†</sup>, Chung-Che Wang, Jyh-Shing Roger Jang “[Adversarial Speaker Distillation for Countermeasure Model on Automatic Speaker Verification](#)”. in *Proc. ISCA SPSC 2022, Incheon, Korea, Sept 2022.*

## Awards

|           |  |                         |
|-----------|--|-------------------------|
| 2024      | <b>Google Student Travel Grant</b> , Google LLC  | <i>Travel Grant</i>     |
| 2024      | <b>CTCI Foundation Bursary Award for Overseas Students</b> , Taiwan CTCI Foundation                      | <i>Bursary Award</i>    |
| 2020-2025 | <b>Distinguished Academic Record Award (5 years)</b> , Taipei Kwong Tong Community Associations          | <i>Scholarship</i>      |
| 2021      | <b>Ranked 3rd of the logical access track of the ASVspoof 2021 challenge</b> , Interspeech 2021          | <i>3rd/42 Worldwide</i> |
| 2018-2020 | <b>Certificate of Achievement (Top 5%, 3 semesters)</b> , NTUST  | <i>Scholarship</i>      |
| 2017      | <b>National Bronze &amp; Guangdong Provincial Gold Awards</b> , 3rd China College Internet Entrep. Comp. | <i>3rd Nationwide</i>   |
| 2016-2017 | <b>National Encouragement Scholarship (Only 3%) &amp; 3rd Prize of Academic Award (Top 20%)</b>          | <i>SZIT Scholarship</i> |

# Research Experiences

Speech Processing and Machine Learning Lab, NTU, Advisor: Prof. [Hung-yi Lee](#)

Sept. 2023 - Present

RESEARCH ASSISTANT

- **[Research]** Researched on **Neural Audio Codec Model** and **Multimodal Large Language Model**
  - Co-built a large-scale spoken language model benchmark, ensuring standardized evaluation tasks. [\[ICLR '25\]](#)
  - Co-built a neural audio codec benchmark [\[ACL Findings '24\]](#) and co-organize a neural audio codec challenge [\[SLT '24\]](#).
  - Contributed to an audio language modeling report comparing methodologies systematically. [\[arXiv '24\]](#)
  - Contributed to a technical report on spoken language models for neural audio codec evaluation. [\[arXiv '24\]](#)
- **[Project]** Leading a student AI team in **TAIHUCAIS: Taiwan Humanities Conversational AI Knowledge Discovery System**
  - Developing a multi-agent system for humanities research scenarios based on Large Language Models [\[Demo\]](#).

Multimedia Information Retrieval Lab, NTU, Advisor: Prof. [Jyh-Shing Roger Jang](#)

Oct. 2020 - Present

RESEARCH ASSISTANT

- **[Research]** Researched on **Multimodal Biometric Authentication**. Co-Advisor: Prof. [Hung-yi Lee](#); Mentor: Dr. [Haibin Wu](#)
  - Proposed a effective & efficient audio-visual active speaker detection model via Transformer distillation [\[ICASSP 2024\]](#)
  - Proposed a SingGraph model using music and speech foundation models for SingFake detection [\[Interspeech 2024\]](#)
  - Proposed neural codec-based adversarial sample detection method for speaker verification [\[Interspeech 2024\]](#)
  - Built a diffusion and flow-matching based deepfake speech dataset [\[SLT 2024\]](#)
  - Analyzed model vulnerabilities and proposed a Push-Pull loss function against adversarial attack [\[SLT 2022\]](#)
  - Proposed an effective and efficient anti-spoofing model via two-stage distillation schema [\[ISCA SPSC 2022\]](#).
  - Ranked **3rd out of 42 submissions** in the logical access track of the ASVspoof 2021 challenge of [Interspeech 2021](#).
- **[Research]** Researched on **Speech Enhancement** and **Source Separation**
  - Proposed a MAMBA-based multichannel speech enhancement model using spectral-spatial information [\[ICASSP 2025\]](#).
  - Built a singer separation dataset for karaoke content generation [\[O-COCOSDA 2024\]](#).

# Services

|           |  |
|-----------|--|
| 2023-Now  | <b>Reviewer / Program Committee</b> , EMNLP (2025), ACL (2024), ICASSP (2023-2025), LREC-COLING/COLING (2024-2025), MLSP (2024), IALP (2024), ECCV AVGenL (2024) |
| 2023-Now  | <b>Administrative Assistant</b> , NVIDIA-NTU Artificial Intelligence Joint Research Center, NTU  |
| 2024      | <b>Invited Talker</b> , Topic: “Singing Voice Graph Modeling for SingFake Detection”, Special Session: SVDD @ IEEE SLT 2024                                      |
| 2024      | <b>Technical Committee</b> , Codec-SUPERB Challenge at IEEE 2024 Spoken Language Technology Workshop (SLT 2024)  |
| 2024-2025 | <b>Teaching Assistant</b> , EE5200: <a href="#">Introduction to Generative AI (2024 Spr.)</a> and EE5184: <a href="#">Machine Learning (2025 Spr.)</a> , NTU     |

# Open Source and Skills

|                         |   |
|-------------------------|---|
| <b>Open Source</b>      | Official repository of MTDVocaLiST model for Audio-Visual Synchronization task <a href="#">(Starred 24)</a> |
| <b>Open Source</b>      | Official repository of SingGraph model for SingFake Detection task <a href="#">(Starred 23, Fork 4)</a>     |
| <b>PLs and Toolkits</b> | Python, LaTeX, Shell; PyTorch, Huggingface, PyTorch Lightning, Numpy, Pandas, Matplotlib.                   |